

$$\frac{1}{1+e^{-z}} = f(z)$$

$$f'(z) = \frac{-1}{(1+e^{-z})^2} \cdot e^{-z} \cdot (-1)$$

$$= \frac{e^{-z}}{(1+e^{-z})^2} = \left(\frac{1}{1+e^{-z}} \right) \left(1 - \frac{1}{1+e^{-z}} \right)$$

$$= f(z)(1-f(z))$$

RMS prob

$$W_{n+1} = W_n - \boxed{\gamma} \frac{\nabla_w L}{|C_{n+1}|}$$

adam

$$\boxed{C_{n+1} = p(C_n) + (1-p) \nabla_w L}$$

momentum.

$$W_{n+1} = W_n - \gamma V_{n+1}$$

$$V_{n+1} = \rho V_n + \eta \nabla_w L \quad \rho = 0.9$$

$$y_1 = \underbrace{W_1^T}_{\text{weights}} \underbrace{x_1}_{\text{input}} + \underbrace{b_1}_{\text{bias}}$$

$$\underbrace{y_2}_n = W_2^T \underbrace{y_1}_m + b_2$$

$$y_n = W_n^T y_{n-1} + b_{n-1}$$

$$y = w^T x + b = \underbrace{\sum_{i=0}^n w_i x_i + b}_{}$$

$$\text{Var}(y) = ?$$

$$\text{Var}(y) = \mathbb{E}(w x)^2 - (\mathbb{E} w x)^2$$

$$= \mathbb{E}(w)^2 \mathbb{E}(x^2) - (\mathbb{E} w \mathbb{E} x)^2$$

$$= (\mathbb{E} w^2 - (\mathbb{E} w)^2 + (\mathbb{E} w)^2) (\mathbb{E} x^2 - (\mathbb{E} x)^2 + (\mathbb{E} x)^2) - (\mathbb{E} w \mathbb{E} x)^2 =$$

$$= (\text{Var } w + (\mathbb{E} w)^2) (\text{Var } x + (\mathbb{E} x)^2) - (\mathbb{E} w \mathbb{E} x)^2$$

$$\begin{aligned}
 &= \text{Var } w \text{ Var } x + \text{Var } x (E w)^2 + \\
 &+ \text{Var } w \cdot (E x)^2 + \cancel{E x^2 E w^2} - \\
 &- \cancel{E x^2 E w^2}
 \end{aligned}$$

$$w \sim N(0, 1)$$

$$\text{Var}(y_i) = \text{Var}(w) \cdot \text{Var}(x)$$

$$\text{Var}(y) = \sum_{i=0}^k \text{Var}(y_i) = k \cdot \text{Var}(y_i)$$

$$= k \text{Var}(w) \text{Var}(x)$$

$$w \sim U\left[-\frac{1}{\sqrt{k}}, \frac{1}{\sqrt{k}}\right]$$

$$\text{Var}(w) = \frac{1}{12} \left(\frac{2}{\sqrt{k}} \right)^2 = \frac{1}{3k}$$

$$\begin{aligned} \text{Var}(y) &= k \cdot \frac{1}{3 \cdot k} \cdot \text{Var}(x) = \\ &= \frac{1}{3} \text{Var}(x). \end{aligned}$$

$$\text{Var}(\hat{y}) = \left(\frac{1}{3} \right)^{100} \cdot \text{Var}(x) \approx 0$$

$$\text{Var}(w) = \frac{2}{k + m}$$

$$W \sim U \left[-\frac{\sqrt{6}}{\sqrt{k+m}}, \frac{\sqrt{6}}{\sqrt{k+m}} \right]$$

[Glorot Uniform]

He.

$$W \sim N(0, \sqrt{\frac{2}{n}})$$